scientific reports

Check for updates

OPEN

Genomic characterization of clinical *Stenotrophomonas* strains from Thailand reveals five putative novel genospecies and extensive functional diversity

Ei Phway Thant[1], Chollachai Klaysubun[1], Prasit Palittapongarnpim[2,3], Kamonnut Singkhamanan[1], Mingkwan Yingkajorn[4], Thunchanok Yaikhan[1], Nattarika Chaichana[1], Rattanaruji Pomwised[5], Monwadee Wonglapsuwan[5], Sarunyou Chusri[6], Arnon Chukamnerd[6], Sissades Tongsima[7], Worawich Phornsiricharoenphant[7], Pimonwan Phokhaphan[7], Sirinthorn Sunthornthummas[7], Nirinya Sudtachat[7], Chumpol Ngamphiw[7] & Komwit Surachat[1,8]✉

Stenotrophomonas spp. are increasingly recognized as opportunistic pathogens with high levels of antimicrobial resistance (AMR). Despite their clinical importance, genomic diversity within the genus remains underexplored, particularly in Thailand and Southeast Asia. Ten *Stenotrophomonas* strains were isolated from diverse clinical specimens at Songklanagarind Hospital, Thailand, in 2023. Whole genome sequencing was performed using MGISEQ 2000 platform. Genome-based classification was assessed via average nucleotide identity (ANI), and digital DNA-DNA hybridization (dDDH). Functional annotation was conducted using RAST and COG databases. Pan-genome structure, AMR/virulence genes, and phenotypic traits (biofilm and hemolysis) were investigated. The genomes ranged from 4.11 to 5.05 Mb with 66.02–66.84% GC content. While 16S rRNA gene sequences showed high similarity to type strains, ANI and dDDH analyses clearly differentiated the ten studied strains from validly published species and grouped them into five putative novel genospecies. The pan-genome was open, with only 21.6% core, 1.5% soft-core and 76.9% accessory/unique genes, indicating high plasticity. Functional annotation revealed enrichment in genes related to regulation, metabolism, and cell envelope biogenesis, reflecting metabolic flexibility and environmental adaptability. AMR profiling showed conserved aminoglycoside-modifying enzymes, β-lactamase bla_{L1} and efflux pumps across all strains. Phenotypically, all strains exhibited multidrug resistance but remained uniformly susceptible to cotrimoxazole, the current treatment of choice. Virulence genes for adhesion, hemolysin, proteases, and biofilm were conserved, consistent with observed α-hemolysis and moderate-to-strong biofilm formation. This study reveals substantial genomic and functional diversity among ten *Stenotrophomonas* strains, highlighting five putative novel genospecies and underscoring the importance of continued surveillance and accurate identification for effective clinical management strategies.

Keywords *Stenotrophomonas* species, ANI, dDDH, Cytotoxic, Biofilm, Antibiotic resistance

[1]Department of Biomedical Sciences and Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand. [2]Department of Microbiology, Faculty of Science, Mahidol University, Bangkok, Thailand. [3]Pornchai Matangkasombut Center for Microbial Genomics, Department of Microbiology, Faculty of Science, Mahidol University, Bangkok, Thailand. [4]Department of Pathology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkla 90110, Thailand. [5]Division of Biological Science, Faculty of Science, Prince of Songkla University, Songkhla 90110, Thailand. [6]Division of Infectious Diseases, Department of Internal Medicine, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand. [7]National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Thailand Science Park, Pathum Thani 12120, Thailand. [8]Translational Medicine Research Center, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand. ✉email: komwit.s@psu.ac.th

The genus *Stenotrophomonas* comprises a diverse group of Gram-negative, aerobic, non-fermenting bacteria that inhabit a wide variety of ecological niches, including soil, plant rhizospheres, freshwater, and clinical environments. Initially described in association with human infection, *S. maltophilia*, one of the main *Stenotrophomonas* species, was originally classified as *Pseudomonas maltophilia*, and later reassigned to the genus *Stenotrophomonas* based on its distinct characteristics[1,2]. The genus currently comprises 27 validly published and correctly named species according to the List of Prokaryotic Names with Standing in Nomenclature (LPSN) database, and genome data are available for 24 of these species in the form of RefSeq representative genome assemblies deposited in the NCBI database as of June 2025[3,4]. Some members of this genus are increasingly recognized for promising biotechnological traits, such as bioremediation, plant growth promotion, or enzyme production, whereas others, such as *S. maltophilia*, *S. sepilia* and *S. pavanii*, have emerged as global opportunistic human pathogens capable of causing respiratory, bloodstream, and urinary infections[5–8].

Clinically relevant species have garnered increasing attention due to their natural resistance to a broad spectrum of antibiotics, attributed to various mechanisms including chromosomally encoded multidrug efflux pumps, production of enzymes, such as aminoglycoside-inactivating enzymes and β-lactamases, and biofilm formation capacity, which complicate treatment options and contribute to poor clinical outcomes[9–13]. Moreover, the ability of members of the genus *Stenotrophomonas* to thrive in hospital environments, including disinfectants, water systems, and medical devices, underscores its role as a formidable opportunistic pathogens, particularly affecting immunocompromised patients[9,14,15]. These characteristics not only highlight the urgent need for effective surveillance and infection control strategies but also emphasize the importance of genomic and taxonomic refinement to track strain-level differences associated with virulence and antimicrobial resistance.

However, traditional methods based on MALDI-TOF or 16S rRNA gene sequencing offer limited resolution for accurate species delineation within *Stenotrophomonas*, especially among closely related taxa within the *S. maltophilia* complex (Smc), a group comprising cryptic species that are not distinguishable by conventional phylogenetic markers[16–18]. In recent years, the implementation of whole-genome-based metrics, such as average nucleotide identity (ANI), digital DNA-DNA hybridization (dDDH), and phylogenomic analysis, has significantly improved taxonomic resolution and led to the reclassification and discovery of novel species[19,20].

In our previous study, we conducted a comprehensive genomic analysis of *Stenotrophomonas* strains isolated from clinical specimens collected from Songklanagarind Hospital in Southern Thailand. Based on species delineation thresholds of ANI (≥95%) and dDDH (≥70%), along with robust phylogenomic evidence, five genetically distinct lineages (hereafter referred to as genospecies) encompassing ten isolates within the genus *Stenotrophomonas* were identified. This study expands the known taxonomic diversity of the genus and provides valuable genomic resources for future ecological and evolutionary research, while also offering critical insights into clinical significance, pathogenic potential, and antimicrobial resistance mechanisms of strains isolated from hospital settings.

Results and discussion

Bacterial isolation and initial identification

The strains PSU_St7, PSU_St15, PSU_St19, PSU_St22, PSU_St103, and PSU_St154 were isolated from sputum samples; PSU_St83 and PSU_St142 from body fluids; PSU_St10 from blood culture; and PSU_St99 from urine. All isolates were obtained in 2023 from the Clinical Microbiology Laboratory at Songklanagarind Hospital, Southern Thailand. Initial phenotypic identification using the MALDI-Biotyper identified all strains as *Stenotrophomonas maltophilia*.

DNA sequencing and functional genome analyses

The de novo assembled genomes varied in size, ranging from approximately 4.11 to 5.05 million base pairs (bp), with an average genome length of about 4.68 million bp. The GC content among the isolates spanned from 66.02 to 66.84% (Table S1). Genome completeness was estimated to be over 99%, indicating high assembly quality of the genomic data.

The functional annotation of the genomes from ten *Stenotrophomonas* strains was performed using RAST SEED and COG classification systems to investigate metabolic capabilities and functional potential (Fig. 1)[21,22]. RAST SEED analysis assigned genes to a broad range of subsystems, with metabolism-related functions consistently dominant across all genomes. Core functional groups included amino acid metabolism, carbohydrate metabolism, the biosynthesis of cofactors, vitamins, prosthetic groups, and pigments, protein metabolism, respiration, lipid metabolism, nitrogen metabolism, and membrane transport, highlighting the

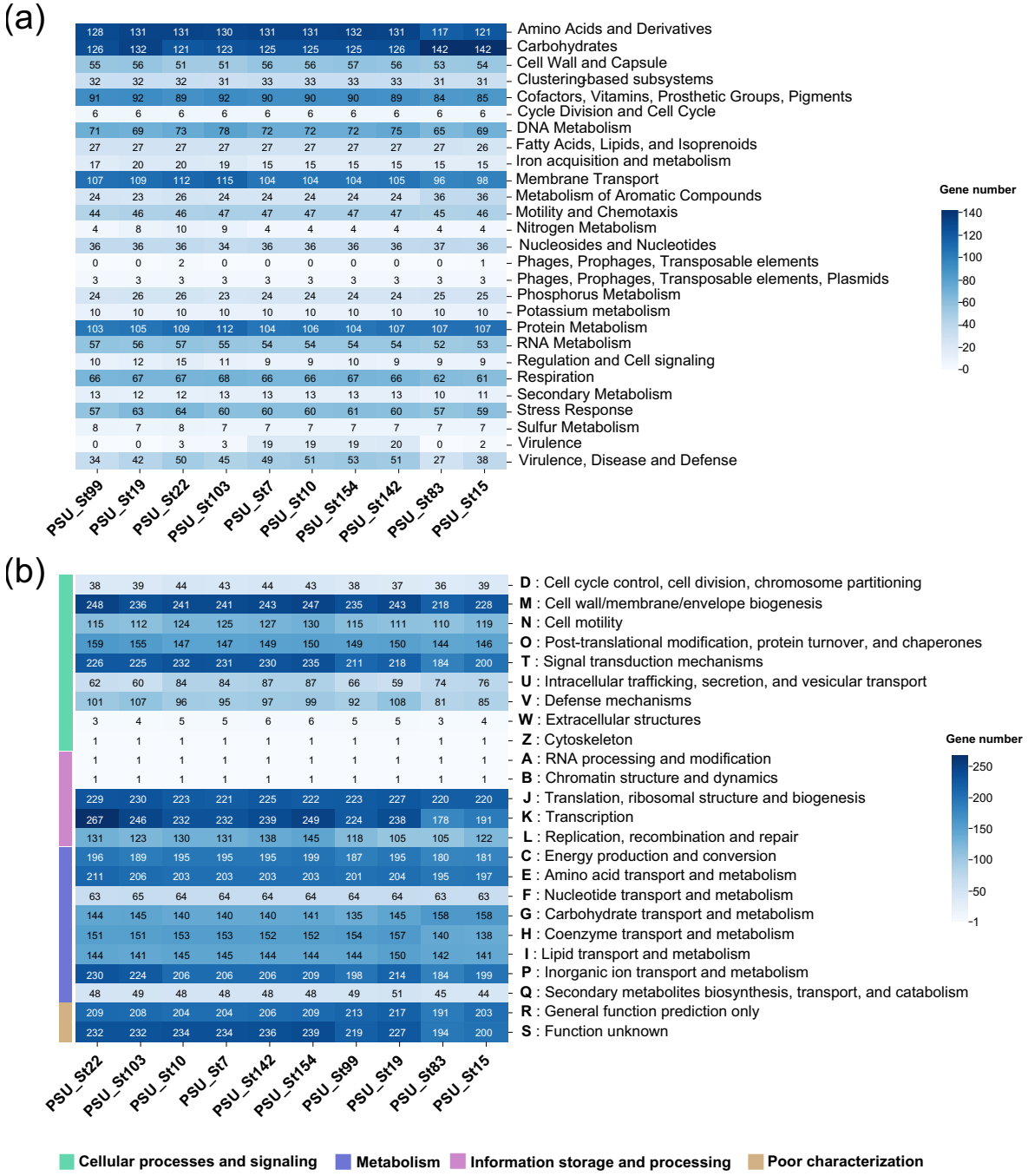


Fig. 1. Functional categorization of annotated genes for ten *Stenotrophomonas* strains. (**a**) Subsystem distribution based on RAST SEED analysis. (**b**) Classification of genes according to COG functional categories. Heatmap intensity corresponds to gene count, with darker shades indicating higher numbers. The figure was generated from data obtained using the KBase "View Function Profile for Genomes–v1.6.0" application and visualized using custom Python scripts.

strains' metabolic versatility. These subsystems are fundamental for cellular maintenance, suggesting that all strains harbor a conserved core of essential metabolic pathways[23].

The COG-based functional classification (Fig. 1b) further supported the functional diversity among the strains. High gene abundance was observed in categories associated with cell wall and cell membrane biogenesis, signal transduction metabolism, translation, ribosomal structure and biogenesis, transcription, amino acid transport and metabolism, carbohydrate transport and metabolism, inorganic ions, transport and metabolism. The consistent presence of genes within these categories across strains suggests a shared core genome architecture, enabling adaptation to diverse ecological or clinical environments[23]. This analysis highlights a genomic framework that is both robust and adaptable, with *Stenotrophomonas* spp. possessing a versatile

toolkit for metabolism, stress responses, and environmental sensing. The substantial proportion of "function unknown" genes within the COG dataset likely represents strain-specific or niche-specific adaptations involving poorly characterized proteins that may enhance environmental specialization, consistent with observations that accessory genomes and particularly unassigned COG functions often underpin microbial adaptation to specific habitats[24]. These comparative functional profiles provide compelling insights into the metabolic versatility and ecological competitiveness of the strains and support their delineation from other members of the *Stenotrophomonas* genus. This functional adaptability may underpin their success in fluctuating environments or under antimicrobial pressure[23].

16S rRNA gene sequences and phylogeny

The complete 16S rRNA gene sequences of the strains PSU_St7, PSU_St15, PSU_St19, PSU_St22, PSU_St103, PSU_St154, PSU_St83, PSU_St142, PSU_St10 and PSU_St99 were retrieved from their genomes. All strains showed more than 99.86% sequence similarity to the closest type strains of validly published *Stenotrophomonas* species, which is higher than the established species differentiation threshold of 98.65%[25]. These results indicate that the 16S rRNA gene sequence has low discriminability for *Stenotrophomona*s species. Phylogenetic analysis based on 16S rRNA gene sequences confirmed that all the strains belonged to the genus *Stenotrophomonas* (Fig. 2, S1-2).

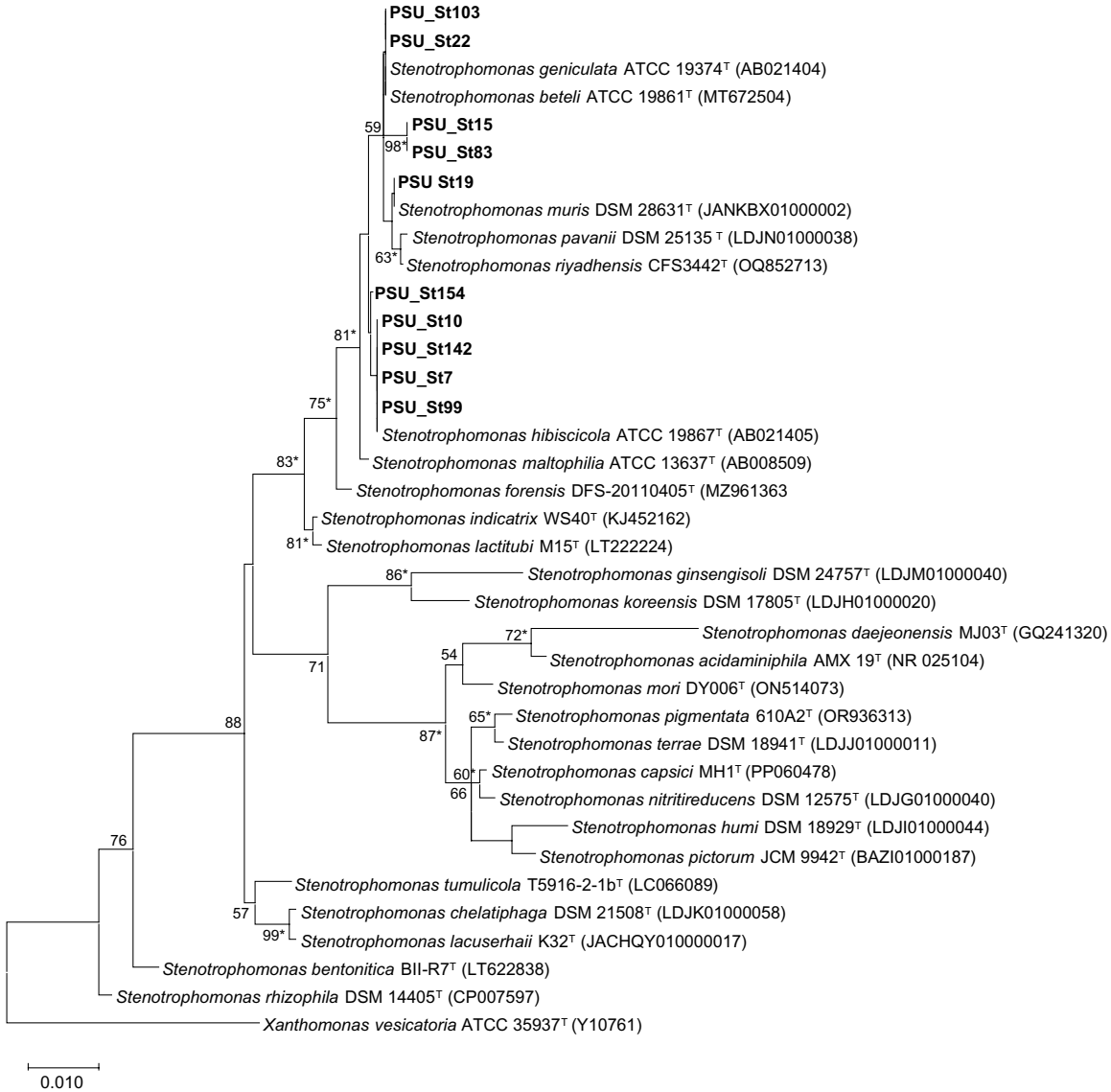


Fig. 2. Phylogenetic analysis of *Stenotrophomonas* strains. Maximum-likelihood tree based on 16S rRNA gene sequences of the strains compared to *Stenotrophomonas* type strains. Bootstrap value (percentage) was computed based on 1,000 bootstrap replicates. The trees were visualized using MEGA X, and bootstrap values above 50% are displayed on the tree. *Xanthomonas vesicatoria* ATCC 35937 ^{T} was employed as an outgroup. Asterisks indicate nodes that were also supported by NJ and MP analyses.

Whole genome-based taxonomic placement and phylogenetic analyses

Average nucleotide identity (ANI) analysis was performed to assess the taxonomic classification of ten *Stenotrophomonas* strains isolated in this study in comparison with type strains of validly published *Stenotrophomonas* species. The resulting heatmap (Fig. 3) displays a pairwise ANI matrix, with darker blue indicating lower similarity and red representing high similarity. Based on the ANI values, the ten isolates demonstrated genomic distinctiveness from all currently recognized *Stenotrophomonas* species, with inter-strain ANI values compared to type strains falling below the 95% species delineation threshold, with the highest ANI value observed between PSU_St22 and a known type strain of *S. riyadhensis* being 93.05%[19]. This finding supports their classification as novel genospecies within the genus *Stenotrophomonas*. Further analysis revealed that the isolates formed five distinct genomic clusters, each with intra-group ANI values exceeding 95%, indicating high genomic similarity among strains within the same cluster.

These findings suggest that the isolates do not belong to any known *Stenotrophomonas* species but rather represent previously undescribed genomic lineages within the genus. The ANI heatmap thus provides strong evidence for the taxonomic novelty and the genetic diversity of these strains.

Phylogenomic analysis via TYGS further supported these delineations. TYGS-based tree inference using Genome BLAST Distance Phylogeny (GBDP) confirmed the placement of the ten PSU strains into five well-separated clusters, with dDDH values well below the 70% species delineation threshold (33.1–52.9%; Fig. 4, S3). Furthermore, a total of 100 single-copy core genes were identified and used to reconstruct a phylogenomic tree using autoMLST (Fig. S4). Based on these results, PSU_St7, PSU_St142, PSU_St10, and PSU_St154 were assigned to genospecies 1; PSU_St99 to genospecies 2; PSU_St22 and PSU_St103 to genospecies 3; PSU_St19 to genospecies 4; and PSU_St15 and PSU_St83 to genospecies 5.

Collectively, these findings indicate that the ten PSU strains represent five distinct genomic lineages within the genus *Stenotrophomonas*. We emphasize that formal species designation requires additional phenotypic, chemotaxonomic, and genomic evidence (e.g., complete genomes generated using long-read sequencing); therefore, these lineages are proposed here as putative novel genospecies. Although the genomes analyzed in this study are draft assemblies based on short-read sequencing, their high completeness supports robust comparative genomic and phylogenomic inference. Future studies incorporating complete genomes and detailed phenotypic characterization will be essential to validate the taxonomic status and clinical significance of these strains.

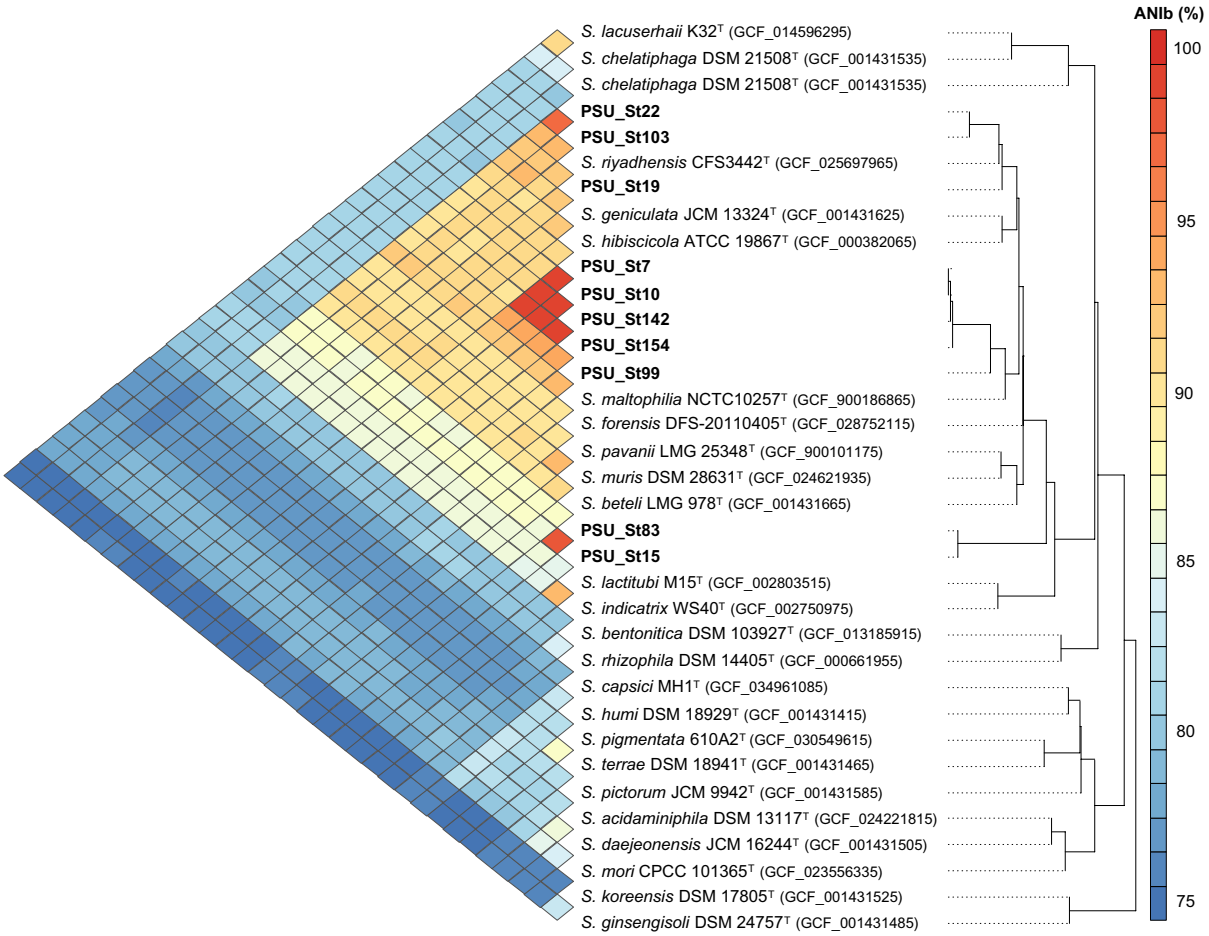


Fig. 3. Average nucleotide identity analysis of ten *Stenotrophomonas* strains isolated in this study in comparison with type strains of validly published *Stenotrophomonas* species.

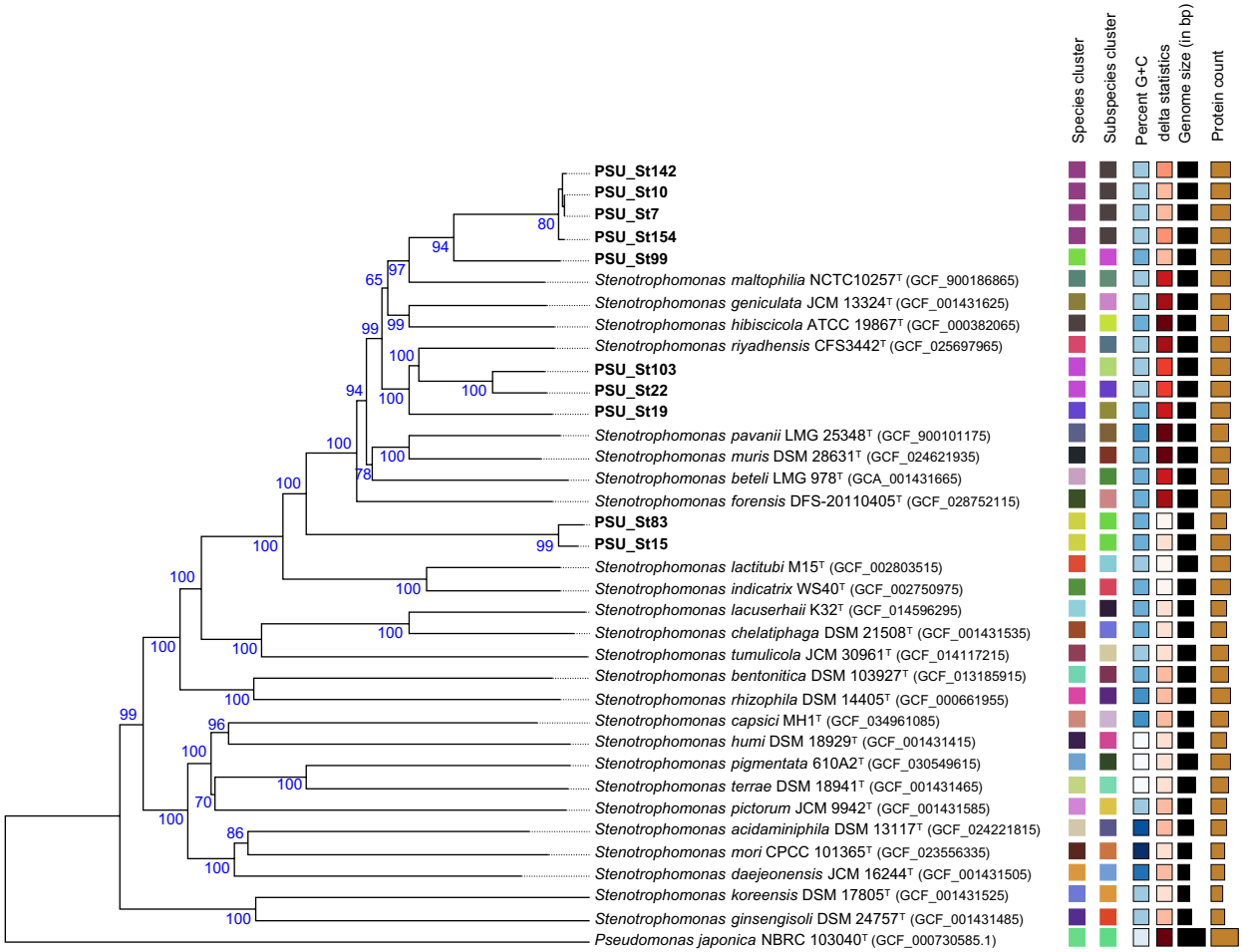


Fig. 4. Phylogenomic analysis based on TYGS showing relationship between the ten strains with closely related species of the genus *Stenotrophomonas*. The outgroup used was *Pseudomonas japonica* NBRC 103040 ^{T}. The tree was constructed with FastME 2.1.6.1 using GBDP distances calculated from genome sequences, with branch lengths reflecting the GBDP distance formula d5. Numbers above branches represent GBDP pseudo-bootstrap support values exceeding 60% from 100 replications, with an average branch support of 92.1%. The tree was rooted at the midpoint.

Pangenome analysis

Pan-genome analysis of the ten studied strains with closely related *Stenotrophomonas* type strains revealed extensive genomic diversity, comprising a total of 11,720 gene clusters. These included core genes (21.62%, n = 2,534), soft-core genes (1.46%, n = 171), shell genes (24.69%, n = 2,894), and cloud genes (52.23%, n = 6,121) (Fig. 5a). A cumulative curve was generated by PanGP based on Heaps' law and is shown in Fig. 5b. According to the γ parameter derived from Heaps' law, the pan-genome is considered to be in an "open" state, with a γ value of 0.46 (γ > 0). This suggests that the species is not saturated, and the inclusion of additional genomes will continue to expand the pan-genome. The predominance of accessory genes (76.9%) further supports the concept of an open pan-genome in *Stenotrophomonas*, likely driven by niche-specific adaptation, horizontal gene transfer, and ecological diversification[26,27].

The gene presence/absence matrix further corroborates these findings, demonstrating clear clustering of the studied strains into distinct genomic groups consistent with core genome phylogeny (Fig. 5c). Specifically, PSU_St7, PSU_St142, PSU_St10, and PSU_St154 (genospecies 1) and PSU_St99 (genospecies 2) exhibited a relatively conserved accessory genome distinct from other groups, whereas PSU_St83 and PSU_St15 (genospecies 5) displayed a more divergent profile, including several unique gene clusters, supporting their classification as a novel lineage. PSU_St103 and PSU_St22 (genospecies 3) formed an intermediate cluster, while PSU_St19 (genospecies 4) emerged as the most genetically divergent strain, characterized by the absence of many shared accessory genes and the presence of unique loci. Collectively, these findings strengthen the proposed genospecies delineation of the studied strains and highlight the remarkable genomic plasticity within the *Stenotrophomonas* genus, underscoring their potential for ecological adaptation and the acquisition of novel functional traits[28].

To better understand the functional roles of core, accessory, and unique gene families, COG functional analysis was conducted on the selected strains. The resulting distribution highlighted four major functional groups: information storage and processing, metabolism, cellular processes and signaling, and a cohort of poorly

characterized genes (Fig. S5). The category of general function prediction appeared consistently across core, accessory, and unique genes (Fig. 5d). Core genes were primarily associated with general function prediction, amino acid transport and metabolism, transcription, and translation, and ribosomal structure and biogenesis. Accessory genes also showed enrichment in general function prediction, along with transcription, cell wall/membrane/envelope biogenesis, and inorganic ion transport and metabolism. In comparison, unique genes were primarily associated with general function prediction, cell wall/membrane/envelope biogenesis, and replication, recombination, and repair.

A substantial proportion of predicted proteins were annotated as hypothetical or uncharacterized. To further interpret these genes, we stratified them by pangenome compartment using Roary. Across the pangenome (12,234 gene clusters), 7,758 clusters were classified as hypothetical/uncharacterized. These were strongly enriched outside the core genome, with only 274 clusters in the core genome (21.27% of core genes), compared to 4,208 clusters in the accessory genome (61.71%) and 3,276 clusters in the unique genome (79.38%). Overall, only 3.53% of hypothetical clusters belonged to the core genome, whereas the majority were distributed in accessory (54.24%) and unique (42.23%) compartments. This pattern indicates that uncharacterized genes are predominantly associated with lineage-specific and strain-specific genomic regions, consistent with roles in niche adaptation and genomic plasticity. Representative accessory and unique hypothetical proteins were further analyzed using similarity and domain-based approaches to infer putative functions using BLASTp searches against the NCBI nr and Swiss-Prot databases, along with conserved domain analysis (Table S2).

Functional characterization of representative hypothetical proteins revealed that a subset of accessory and unique genes could be linked to biologically relevant categories, including membrane-associated transport systems, transcriptional regulators, and resistance-related proteins. Notably, several clusters showed similarity to efflux systems (e.g., heavy metal transporters such as CzcB) and regulatory proteins (HTH-type transcriptional regulators), suggesting roles in environmental sensing and adaptive gene regulation. In addition, putative transport and binding proteins were identified, supporting functions in nutrient acquisition and stress tolerance. However, a large proportion of hypothetical proteins lacked significant homology to known proteins, particularly within the unique genome fraction, indicating the presence of lineage-specific or novel gene families. Collectively, these findings suggest that the uncharacterized gene repertoire is not random, but is structured within the accessory and unique genome compartments and likely contributes to ecological versatility, environmental adaptation, and potential opportunistic pathogenicity in *Stenotrophomonas*.

Antimicrobial resistant gene analysis

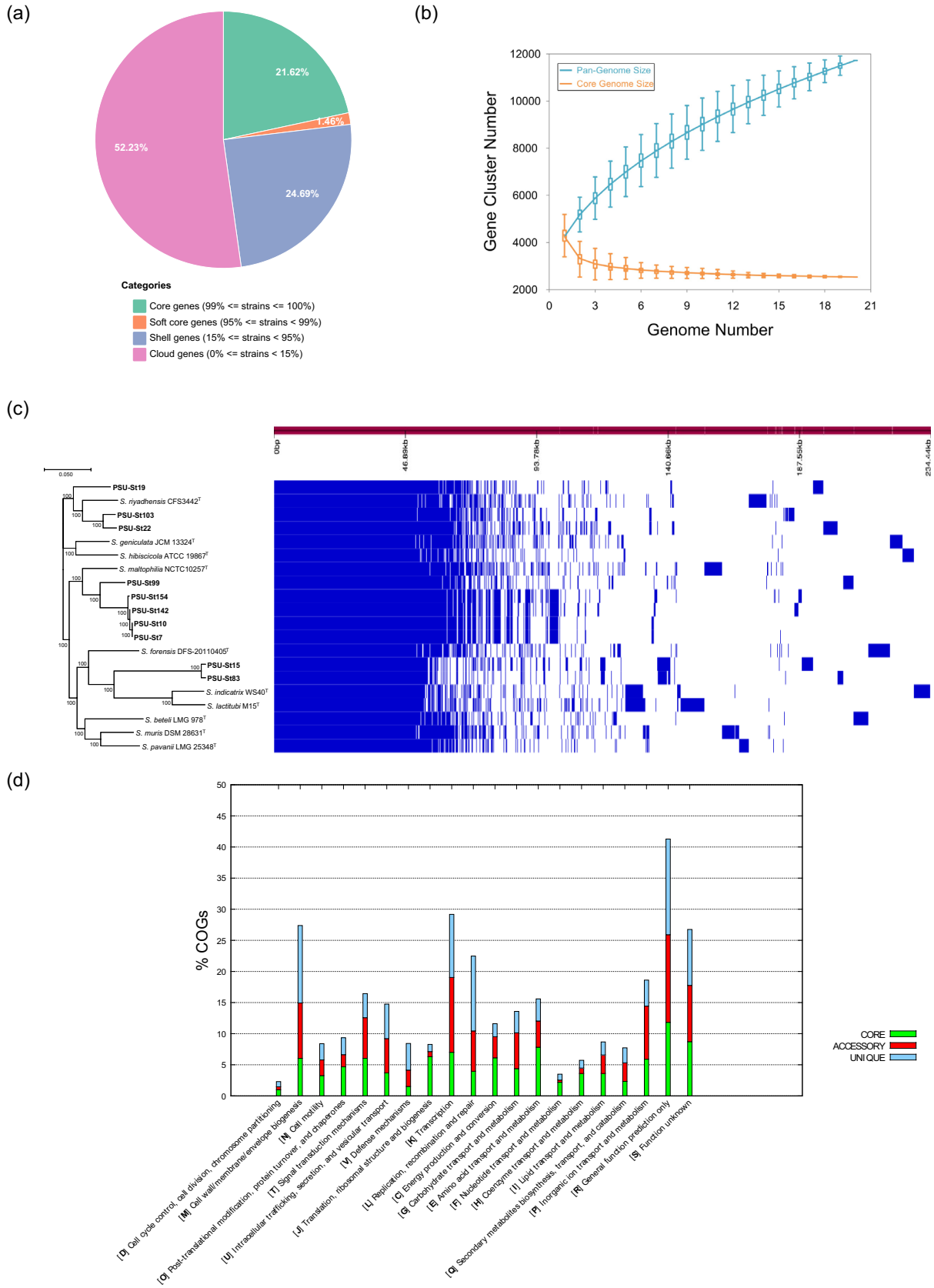
Genospecies 1 (PSU_St7, PSU_St10, PSU_St142, PSU_St154) and genospecies 2 (PSU_St99), which were phylogenetically close to *S. maltophilia* NCTC10257^T, as well as another cluster comprising genospecies 3 (PSU_St22, PSU_St103) and genospecies 4 (PSU_St19), exhibited highly similar resistomes. These included multiple aminoglycoside-modifying enzyme genes (*aph(3′)-IIc*, *aph(9)-Ic*) and the β-lactamase gene bla_{L1}, indicating adaptation to substantial antibiotic pressure, consistent with previous reports on the high prevalence of aminoglycoside resistance genes in the *Stenotrophomonas* genus[26,29]. In contrast, genospecies 5 (PSU_St15, PSU_St83) formed a genetically distinct cluster characterized by a divergent antimicrobial resistance (AMR) gene profile, particularly with respect to aminoglycoside resistance. Efflux pump genes (*smeABC*, *smeDEF*, *smeRS*) were highly conserved across nearly all strains. Prior studies have demonstrated that overexpression of *smeDEF* and *smeABC* contributes to resistance against fluoroquinolones, tetracyclines, and meropenem, supporting their roles in both intrinsic resistance and environmental adaptation, particularly in potential novel *Stenotrophomonas* species[25,26]. These results are shown in Fig. 6.

Phenotypic testing using the agar disc diffusion method revealed consistently high resistance levels to β-lactams, aminoglycosides, and carbapenems across all strains, corroborating the genomic predictions and in line with previous observations of widespread aminoglycoside resistance in the *Stenotrophomonas* genus[26] (Fig. 7). Notably, variable resistance to colistin and ciprofloxacin was observed, particularly in PSU_St7, which remained resistant. This suggests that susceptibility may be influenced by efflux pump expression levels or alternative resistance mechanisms[30]. All studied strains exhibited broad-spectrum phenotypic resistance by antimicrobial susceptibility testing using the disk diffusion method, underscoring their potential clinical relevance and showing strong concordance with genomic resistance predictions. One strain, PSU_St83, which demonstrated intermediate susceptibility to cotrimoxazole—the first-line treatment for *Stenotrophomonas* infections as recommended by the Infectious Diseases Society of America (IDSA) guidelines[31], was further evaluated using the minimum inhibitory concentration (MIC) assay to validate the disc diffusion findings. The MIC result indicated susceptibility (trimethoprim ≤ 1 μg/mL and sulfamethoxazole ≤ 19 μg/mL), and all strains were ultimately found to be susceptible to cotrimoxazole. These findings support the continued clinical efficacy of cotrimoxazole against the proposed novel *Stenotrophomonas* species.

The studied strains were isolated from diverse clinical sources including sputum, urine, and body fluid, whereas closely related type strains originated from various clinical, plant-associated, and environmental sources. These ecological differences may reflect niche-specific adaptations and distinct evolutionary pressures[32]. In conclusion, these findings underscore the taxonomic and clinical relevance of genome-based surveillance and emphasize the necessity of targeting both horizontally acquired resistance genes and intrinsic mechanisms, such as efflux systems, in future antimicrobial strategies.

Virulence factor analysis

Analysis of virulence determinants revealed that the potential novel *Stenotrophomonas* species and closely related type strains share a consistent set of core virulence genes, primarily associated with cytotoxic morphological effects, adhesion, biofilm formation, biofilm maturation, and quorum sensing (Fig. 8). Notably, major extracellular proteases including StmPr1, StmPr2, and StmPr3 were identified, along with additional enzymes



such as hemolysin, DNase, phospholipases, esterases, and alginate lyases, indicating potential for host tissue degradation and evasion of host immune responses[33,34]. Genotypic evidence for the *hly* gene was consistent with phenotypic expression of α-hemolysis observed on 5% blood agar, further validating the genomic prediction. This concordance between genotypic and phenotypic data underscores the functional relevance of hemolysin production in these strains. The presence of these extracellular enzymes in the studied strains, mirroring those

◀**Fig. 5**. Pan-genome analysis of potential novel *Stenotrophomonas* strains and closely *Stenotrophomonas* species. (**a**) The pie chart represents gene distribution across the pan-genome, showing proportions of core, soft core, shell, and cloud genes. (**b**) The rarefaction curves compare pan-gene clusters (blue) and core gene clusters (orange). (**c**) The gene presence/absence matrix (right) visualizes clustering patterns among strains, corresponding to the core genome phylogenetic tree (left). (**d**) COGs distribution of core, accessory, and unique genes of selected strains.

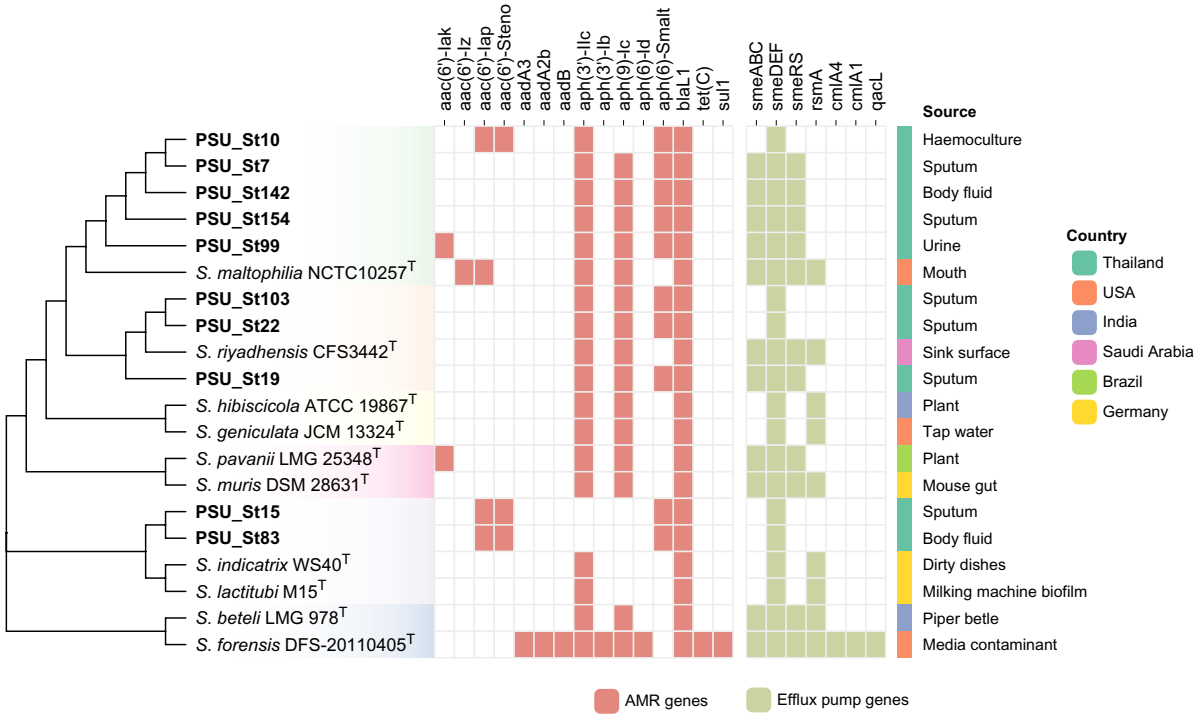


Fig. 6. Antimicrobial resistance gene profiles of potential novel *Stenotrophomonas* strains and closely related reference *Stenotrophomonas* species. Heatmap indicates acquired AMR genes (brick red = AMR genes present) and efflux pump genes (Green = efflux pump genes present) (White = genes absent). Strain sources and countries of origin are color-coded.

found in *S. maltophilia*, highlights the importance of careful monitoring and characterization of these novel species, particularly in clinical or immuno-compromised contexts, where opportunistic infections could occur[6]. The detection of the *narG* gene in a substantial number of isolates suggests an adaptive metabolic capability, potentially enabling anaerobic respiration via nitrate reduction during nutrient limitation. This metabolic versatility may contribute to environmental persistence and colonization in diverse ecological niches[5,34].

Phenotypic biofilm formation assays demonstrated that all studied strains exhibited moderate to strong biofilm production, consistent with the widespread presence of conserved biofilm-associated regulatory and structural determinants across the genomes (Fig. 8 and Table S3). To further investigate the genetic basis of this trait, the presence and absence of biofilm-associated genes were compared with phenotypic biofilm strength. Core genetic determinants involved in surface attachment, quorum sensing, and biofilm maturation were widely conserved, indicating that biofilm development is a fundamental characteristic within the genus *Stenotrophomonas*[35,36]. However, variation in accessory gene content was observed across genospecies, suggesting lineage-specific differences in biofilm-associated mechanisms. Strains exhibiting stronger biofilm production generally possessed a broader repertoire of adhesion- and regulatory-related genes, whereas strains with moderate biofilm levels displayed a comparatively reduced gene complement. Nevertheless, no strict one-to-one correlation between gene presence and phenotype was observed, indicating that additional regulatory mechanisms and environmental factors likely influence biofilm development. Such phenotypic plasticity may enable strains to adapt to diverse ecological or host-associated niches. Understanding the genetic basis of biofilm formation in *Stenotrophomonas* is critical for the development of targeted interventions, such as quorum sensing inhibitors or biofilm-dispersing agents, which may enhance the efficacy of antimicrobial therapies[37,38]. Importantly, this integrated analysis of genotypic and phenotypic biofilm traits provides additional evidence supporting functional diversification among the proposed genospecies.

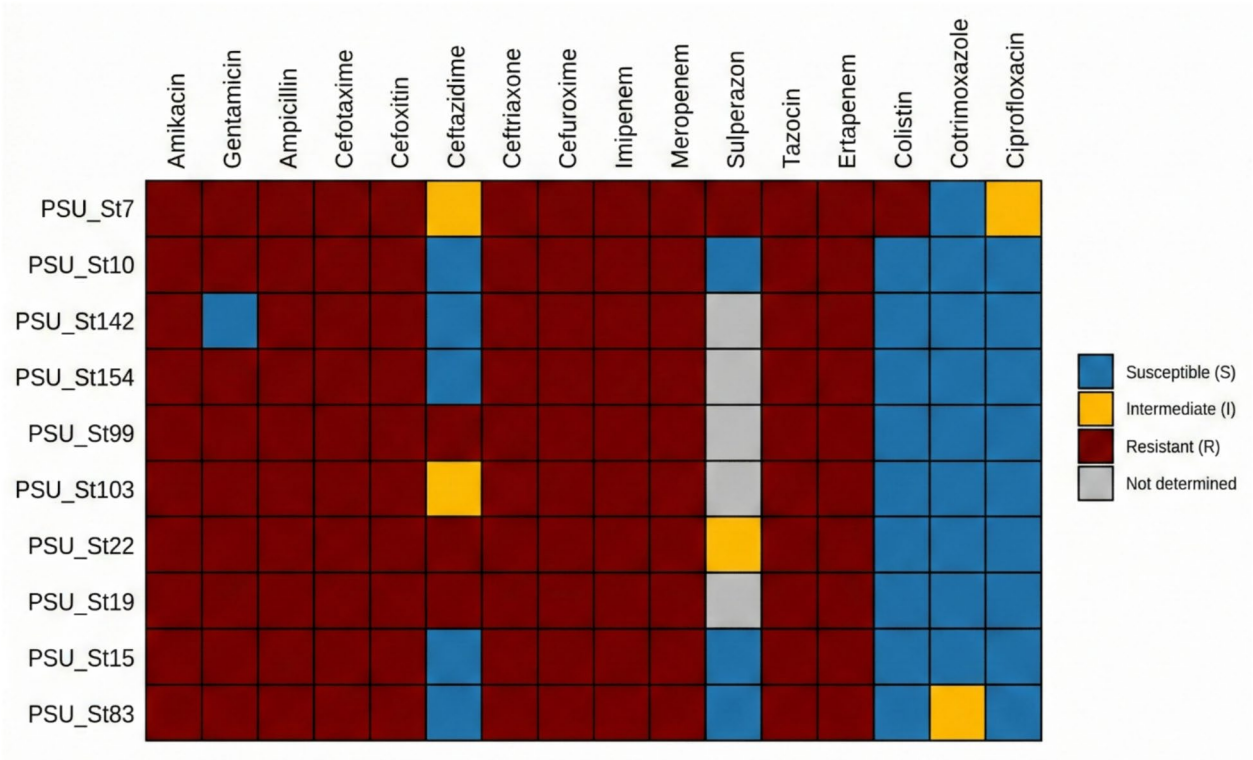


Fig. 7. Phenotypic resistance profiles of novel *Stenotrophomonas* strains to antibiotics. Antibiotic resistance ability is indicated by colored squares as follows: red, resistant; yellow, intermediate; and blue, susceptible.

Conclusion

This study provides a comparative genomic and functional characterization of ten *Stenotrophomonas* isolates, revealing substantial genetic diversity and supporting the presence of five putative novel genospecies. The isolates exhibited diverse multidrug resistance profiles and harbored genes associated with virulence, stress response, and environmental adaptation, highlighting the ecological versatility and potential clinical relevance of the genus. However, this study is limited by the relatively small number of isolates (n = 10), which may not fully capture the genomic diversity of *Stenotrophomonas* in clinical settings. In addition, the absence of in vivo or clinical validation restricts the direct interpretation of pathogenic potential. Future studies incorporating larger, geographically and temporally diverse datasets, as well as functional and infection model validation, will be essential to confirm species boundaries and to further elucidate the clinical significance of these emerging lineages. These findings should be interpreted with caution due to the limited sample size and reliance on draft genome assemblies.

Materials and methods

Bacterial isolation and initial identification

The strains PSU_St7, PSU_St10, PSU_St15, PSU_St19, PSU_St22, PSU_St83, PSU_St99, PSU_St103, PSU_St142, and PSU_St154 were identified using the MALDI-Biotyper system (Bruker Daltonics GmbH, Germany) and collected from the clinical microbiology laboratory at Songklanagarind Hospital, Southern Thailand, during a one-year period in 2023. The isolates were stored in Tryptone Soya Broth (TSB; HiMedia, Nashik, India) supplemented with 20% glycerol at − 80 °C. Relevant clinical and laboratory data, including antibiotic susceptibility profiles of the patients from whom the isolates were obtained, were retrieved and reviewed. *Staphylococcus aureus* ATCC 29,213, acquired from the American Type Culture Collection (ATCC), was used as the reference strain.

Minimum inhibitory concentration assay

The minimum inhibitory concentration (MIC) was tested using the microbroth dilution method, following the Clinical Laboratory Standards Institute (CLSI) guidelines (2020), for the cotrimoxazole-resistant strain identified by the agar disc diffusion method[39].

Determination of biofilm production activity

Biofilm formation was evaluated following the method described by Zhuo et al.[40] with minor modifications. Overnight cultures of *Stenotrophomonas* spp. grown in Tryptone Soya Broth (TSB) were adjusted to a final concentration of 1×10^7 CFU/mL. A volume of 200 µL of the bacterial suspension was added to each well of a sterile, flat-bottom 96-well polystyrene microtiter plate and incubated aerobically at 37 °C for 24 h. Wells

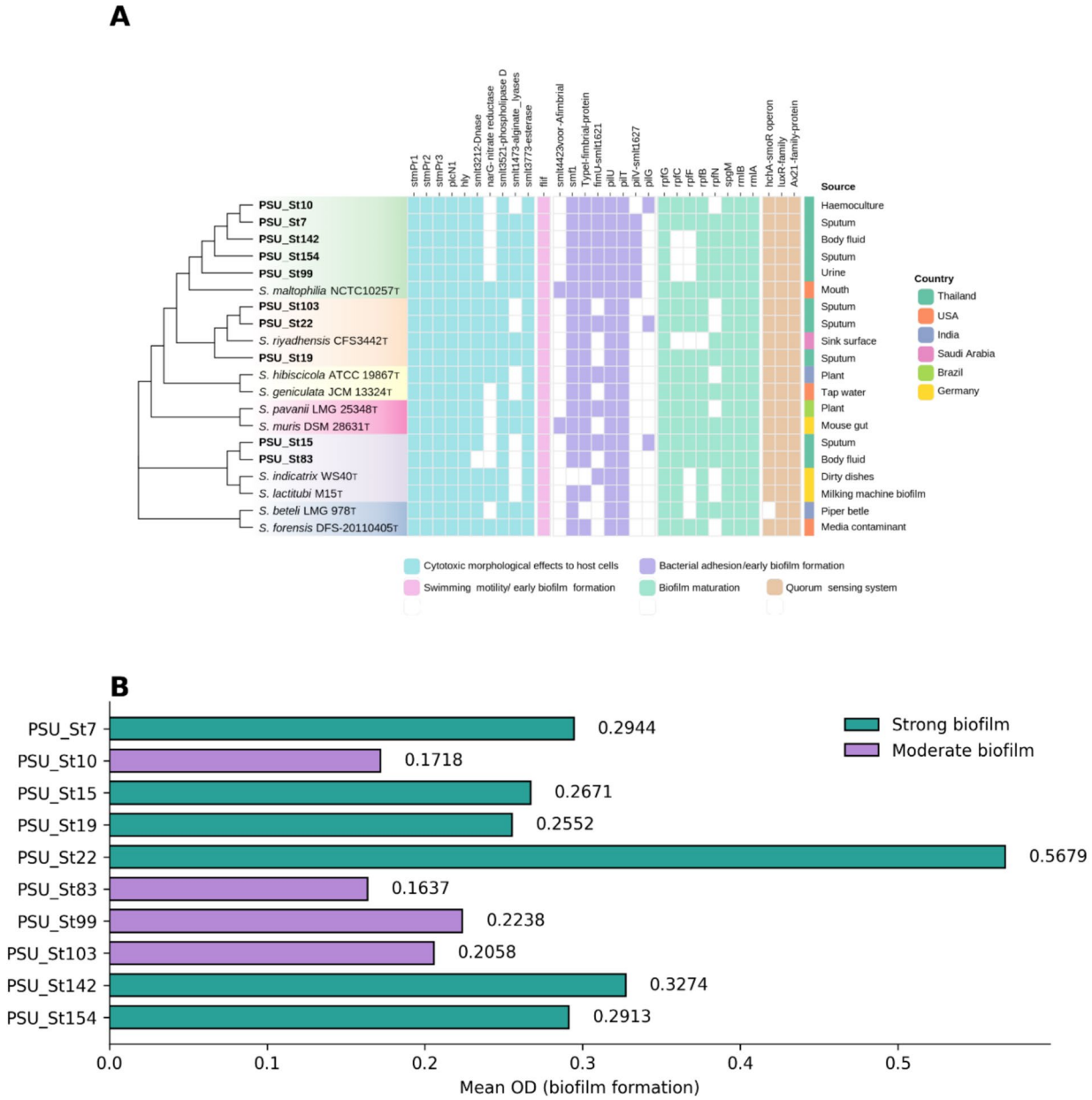


Fig. 8. Virulence factor profiles and biofilm formation in *Stenotrophomonas* strains. (**A**) Presence and absence of virulence-associated genes, including those involved in cytotoxicity, motility, adhesion, biofilm formation, biofilm maturation, and quorum sensing. Colored cells indicate gene presence, while blank cells indicate absence. (**B**) Quantitative biofilm formation measured as mean optical density (OD). Strains were classified as moderate or strong biofilm producers based on predefined OD thresholds (ODc = 0.059; 2× ODc = 0.118; 4× ODc = 0.236). The integration of genotypic and phenotypic data highlights variability in biofilm-associated traits across strains.

containing only the culture medium served as negative controls. After incubation, non-adherent cells were removed by washing twice with 200 µL of sterile phosphate-buffered saline (PBS; pH 7.3; HiMedia, Nashik, India). The plates were then fixed at 60 °C for one hour, stained with 200 µL of 0.1% (w/v) crystal violet (Sigma–Aldrich, St. Louis, USA) for 15 min, and rinsed with tap water to remove excess stain. After air-drying, the bound dye was solubilized by adding 200 µL of 33% glacial acetic acid and incubating at room temperature for 15 min. The optical density was measured at 492 nm (OD_{492}) to quantify the total biofilm biomass. Based on the OD_{492} values, isolates were classified as follows: non-biofilm producers ($OD \leq ODc$), weak ($ODc < OD \leq 2 \times ODc$), moderate ($2 \times ODc < OD \leq 4 \times ODc$), or strong biofilm producers ($OD > 4 \times ODc$), where ODc was defined as the cut-off value calculated as the mean OD of the negative control plus three standard deviations (ODc = mean OD + 3 × SD).

Determination of hemolysis activity

Hemolytic activity was assessed by streaking bacterial isolates onto blood agar plates (HiMedia, Nashik, India) supplemented with 5% sterile defibrinated blood, followed by incubation at 37 °C for 48 h. Hemolysis was evaluated based on the appearance of clear (β-hemolysis), greenish (α-hemolysis), or no (γ-hemolysis) zones surrounding the colonies. *Staphylococcus aureus* ATCC 29,213 was included as a positive control to validate the assay.

DNA extraction and whole-genome sequencing

DNA extraction was performed using the QIAamp® DNA Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. The extracted DNA had a concentration of at least 50 ng/μL and an OD260/280 ratio between 1.8 and 2. DNA integrity was confirmed by 1.5% agarose gel electrophoresis. Whole genome short-read sequencing (150 bp paired-end reads) was conducted at the Beijing Genomics Institute in China using the MGISEQ 2000 platform, following standard protocols.

Genome analysis

Sequence reads were de novo assembled using the SPAdes 3.11.1[41]. The assembled genome quality and completeness were checked by QUAST[42] and BUSCO v5.4.3[43]. The assembled genome was subjected to Prokka and functional annotations on Rapid Annotation using the Subsystem Technology (RAST) server[21,44]. Functional categories based on the COG (Clusters of Orthologous Groups) classification were obtained using the online version of eggNOG-mapper v.2[45].

The phylogenetic tree was constructed by Phandango application (http://phandango.net) using pan-genome analysis data of Roary v3.13.0 to analyze the relationships among *Stenotrophomonas* species[46,47]. The pan-genome was modeled using PanGP[48]. The functional classification based on the COG categorization of the core, accessory, and unique genes was processed and visualized with the BPGA pipeline[49]. The data for comparative genome analysis were retrieved from NCBI Reference Sequence Database[4]. The resulting phylogenetic tree in newick format was further visualized and annotated using Chiplot[50]. Average nucleotide identity (ANI) value was calculated against the downloaded genomic sequences of *Stenotrophomonas* species from NCBI database by using JSpecieWS platform[51]. Digital DNA-DNA hybridization (dDDH) values for putative novel species with low ANI values (below 95%) were analyzed using the Type Strain Genome Server (TYGS) and the Genome-to-Genome Distance Calculator (GGDC)[20,52,53]. Antimicrobial resistance genes were predicted using the ResFinder database (80% identity threshold and 60% minimum length), CARD's RGI database (80% identity threshold with the parameter-perfect, strict, and loose hits) and Abricate database (80% identity threshold) values[54–56]. Whereas IS elements were detected using the ISfinder[57], plasmid identification was initially conducted using Plasme v1.0, which flagged potential plasmid-containing contigs[58]. Confirmation of these plasmids was performed by BLAST analysis against the NCBI database using 99% coverage and identity cut-off[4]. Integrons were identified by using IntegronFinder 2.0[59]. Then, virulence-associated genes were detected by using cut-off values of 80% identity and 1e-30 E-value cut-offs against the virulence factor database (VFDB)[60]. Phage regions were produced using Phigaro2.3.0[61]. Functional profile data were retrieved using the "View Function Profile for Genomes—v1.6.0" application on the KBase platform, and the final comparative visualization was generated using custom Python scripts.

Data availability

The genome assemblies generated in this study have been deposited in the NCBI database under BioProject accession number **PRJNA1399037** . Individual genome assemblies are available under the following accession numbers: GCF_054474035.1, GCF_054473825.1, GCF_054473975.1, GCF_054473875.1, GCF_054473915.1, GCF_054473815.1, GCF_054474005.1, GCF_054473935.1, GCF_054473995.1, and GCF_054473855.1.

Received: 30 December 2025; Accepted: 29 March 2026
Published online: 03 April 2026

References

1. Palleroni, N. J. & Bradbury, J. F. *Stenotrophomonas*, a new bacterial genus for *Xanthomonas maltophilia* (Hugh 1980) Swings et al. 1983. *Int. J. Syst. Bacteriol.* **43**, 606–609 (1993).
2. Swings, J., De Vos, P., Van den Mooter, M. & De Ley, J. Transfer of *Pseudomonas maltophilia* Hugh 1981 to the genus *Xanthomonas* as *Xanthomonas maltophilia* (Hugh 1981) comb. nov.. *Int. J. Syst. Evol. Microbiol.* **33**, 409–413 (1983).
3. Euzéby, J. List of prokaryotic names with standing in nomenclature. http://www.bacterio.cict.fr (2009).
4. Sayers, E. W. et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res* **50**, D20-d26 (2022).
5. Ryan, R. P. et al. The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nat. Rev. Microbiol.* **7**, 514–525 (2009).
6. Brooke, J. S. *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clin. Microbiol. Rev.* **25**, 2–41 (2012).
7. Gautam, V. et al. Description of *Stenotrophomonas sepilia* sp. nov., isolated from blood culture of a hospitalized patient as a new member of *Stenotrophomonas maltophilia* complex. *New Microbes and New Infections* https://doi.org/10.1016/j.nmni.2021.100920 (2021).
8. Fluit, A. C. et al. Taxonomic position, antibiotic resistance and virulence factor production by *Stenotrophomonas* isolates from patients with cystic fibrosis and other chronic respiratory infections. *BMC Microbiol.* **22**, 1–18 (2022).
9. Paopradit, P., Srinitiwarawong, K., Ingviya, N., Singkhamanan, K. & Vuddhakul, V. Distribution and characterization of *Stenotrophomonas maltophilia* isolates from environmental and clinical samples in Thailand. *J. Hosp. Infect.* **97**, 185–191 (2017).
10. Blanca Sánchez, M., Luis Martínez, J. Overexpression of the efflux pumps SmeVWX and SmeDEF is a major cause of resistance to co-trimoxazole in *Stenotrophomonas maltophilia*. (2018).
11. Sánchez, M. B. & Martínez, J. L. SmQnr contributes to intrinsic resistance to quinolones in *Stenotrophomonas maltophilia*. *Antimicrob. Agents Chemother.* **54**, 580–581 (2010).

12. Furlan, J. P. R., Sanchez, D. G., Gallo, I. F. L. & Stehling, E. G. Characterization of acquired antimicrobial resistance genes in environmental *Stenotrophomonas maltophilia* isolates from Brazil. *Microb. Drug Resist.* **25**, 475–479 (2019).
13. Mojica, M. F. et al. Population structure, molecular epidemiology, and beta-lactamase diversity among *Stenotrophomonas maltophilia* isolates in the United States. *MBio* **10**, 10.1128/mbio.00405-00419 (2019).
14. Turrientes, M. C. et al. Polymorphic mutation frequencies of clinical and environmental *Stenotrophomonas maltophilia* populations. *Appl. Environ. Microbiol.* **76**, 1746–1758 (2010).
15. Mojica, M. F., Humphries, R., Lipuma, J. J., Mathers, A. J., Rao, G. G., Shelburne, S. A., Fouts, D. E., Van Duin, D., Bonomo, R. A. Clinical challenges treating *Stenotrophomonas maltophilia* infections: An update. JAC-Antimicrobial Resistance 2022, 4.
16. Mercier-Darty, M. et al. Comparative whole-genome phylogeny of animal, environmental, and human strains confirms the genogroup organization and diversity of the *Stenotrophomonas maltophilia* complex. *Appl. Environ. Microbiol.* **86**, e02919–02919 (2020).
17. Patil, P. P., Kumar, S., Midha, S., Gautam, V. & Patil, P. B. Taxonogenomics reveal multiple novel genomospecies associated with clinical isolates of *Stenotrophomonas maltophilia*. *Microb. Genom.* **4**(8), e000207 (2018).
18. Li, K. et al. *Stenotrophomonas maltophilia* complex: Insights into evolutionary relationships, global distribution and pathogenicity. *Front. Cell. Infect. Microbiol.* **13**, 1325379 (2024).
19. Ciufo, S. et al. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* **68**, 2386–2392 (2018).
20. Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.-P. & Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 1–14 (2013).
21. Aziz, R. K. et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* **9**(1), 75 (2008).
22. Cantalapiedra, C. P., Hernandez-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
23. Zhao, Y., Ding, W.-J., Xu, L. & Sun, J.-Q. A comprehensive comparative genomic analysis revealed that plant growth promoting traits are ubiquitous in strains of *Stenotrophomonas*. *Front. Microbiol.* **15**, 1395477 (2024).
24. Prabha, R., Singh, D. P., Somvanshi, P. & Rai, A. Functional profiling of cyanobacterial genomes and its role in ecological adaptations. *Genomics Data* **9**, 89–94 (2016).
25. Kim, M., Oh, H.-S., Park, S.-C. & Chun, J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **64**, 346–351 (2014).
26. Xu, Y., Cheng, T., Rao, Q., Zhang, S. & Ma, Y. Comparative genomic analysis of *Stenotrophomonas maltophilia* unravels their genetic variations and versatility trait. *J. Appl. Genet.* **64**, 351–360 (2023).
27. Li, Y., Sun, Z.-Z., Rong, J.-C. & Xie, B.-B. Comparative genomics reveals broad genetic diversity, extensive recombination and nascent ecological adaptation in *Micrococcus luteus*. *BMC Genomics* **22**, 1–14 (2021).
28. Tettelin, H. et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. USA* **102**, 13950–13955 (2005).
29. Gajdács, M. & Urbán, E. Prevalence and antibiotic resistance of *Stenotrophomonas maltophilia* in respiratory tract samples: A 10-year epidemiological snapshot. *Health Serv. Res. Manag. Epidemiol.* **6**, 233339281987077–233339281987077 (2019).
30. Alonso, A. & Martinez, J. L. Expression of multidrug efflux pump SmeDEF by clinical isolates of *Stenotrophomonas maltophilia*. *Antimicrob. Agents Chemother.* **45**, 1879–1881 (2001).
31. Tamma, P. D. et al. Infectious Diseases Society of America guidance on the treatment of AmpC β-lactamase-producing enterobacterales, carbapenem-resistant *Acinetobacter baumannii*, and *Stenotrophomonas maltophilia* infections. *Clin. Infect. Dis.* **74**, 2089–2114 (2022).
32. Youenou, B. et al. Comparative genomics of environmental and clinical *Stenotrophomonas maltophilia* strains with different antibiotic resistance profiles. *Genome Biol. Evol.* **7**, 2484–2505 (2015).
33. DuMont, A. L. & Cianciotto, N. P. *Stenotrophomonas maltophilia* serine protease StmPr1 induces matrilysis, anoikis, and protease-activated receptor 2 activation in human lung epithelial cells. *Infect Immun* **85**(12), 10–128 (2017).
34. Crossman, L. C. et al. The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. *Genome Biol.* **9**(4), 174 (2008).
35. Azimi, A., Aslanimehr, M., Yaseri, M., Shadkam, M. & Douraghi, M. Distribution of *smf-1*, *rmlA*, *spgM* and *rpfF* genes among *Stenotrophomonas maltophilia* isolates in relation to biofilm-forming capacity. *J. Global Antimicrob. Resist.* **23**, 321–326 (2020).
36. García, G., Girón, J. A., Yañez, J. A. & Cedillo, M. L. *Stenotrophomonas maltophilia* and its ability to form biofilms. *Microbiol. Res.* **14**, 1–20 (2022).
37. Rabin, N. et al. Biofilm formation mechanisms and targets for developing antibiofilm agents. *Future Med. Chem.* **7**, 493–512 (2015).
38. Mahdhi, A. et al. Use of extracellular polysaccharides, secreted by Lactobacillus plantarum and Bacillus spp., as reducing indole production agents to control biofilm formation and efflux pumps inhibitor in Escherichia coli. *Microb. Pathog.* **125**, 448–453 (2018).
39. Clinical; Institute, L.S. Performance standards for antimicrobial susceptibility testing. (2020).
40. Zhuo, C., Zhao, Q. Y. & Xiao, S. N. The impact of *spgM*, *rpfF, rmlA* gene distribution on biofilm formation in *Stenotrophomonas maltophilia*. *PLoS ONE* **9**, e108409–e108409 (2014).
41. Bankevich, A. et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
42. Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D. & Gurevich, A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34**, i142–i150 (2018).
43. Manni, M., Berkeley, M. R., Seppey, M. & Zdobnov, E. M. BUSCO: Assessing genomic data quality and beyond. *Curr. Prot.* **1**(12), 1323 (2021).
44. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
45. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
46. Page, A. J. et al. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
47. Hadfield, J. et al. Phandango: An interactive viewer for bacterial population genomics. *Bioinformatics* **34**, 292–293 (2018).
48. Zhao, Y. et al. PanGP: A tool for quickly analyzing bacterial pan-genome profile. *Bioinformatics* **30**, 1297–1299 (2014).
49. Chaudhari, N. M., Gupta, V. K. & Dutta, C. BPGA-An ultra-fast pan-genome analysis pipeline. *Sci. Rep.* **6**, 24373 (2016).
50. Xie, J. et al. Tree Visualization By One Table (tvBOT): A web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res.* **51**, W587–W592. https://doi.org/10.1093/nar/gkad359 (2023).
51. Richter, M., Rosselló-Móra, R., Oliver Glöckner, F. & Peplies, J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**, 929–931 (2016).
52. Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Göker, M. TYGS and LPSN: A database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* **50**, D801–D807 (2022).
53. Meier-Kolthoff, J. P. & Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182 (2019).

54. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder–An open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb. Genomics* **8**, 000748 (2022).
55. McArthur, A. G. et al. The comprehensive antibiotic resistance database. *Antimicrob. Agents Chemother.* **57**, 3348–3357 (2013).
56. Seemann, T. Abricate, Github, https://github.com/tseemann/abricate.
57. Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. & Chandler, M. ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **34**, D32-36 (2006).
58. Tang, X., Shang, J., Ji, Y. & Sun, Y. PLASMe: A tool to identify plasmid contigs from short-read assemblies using transformer. *Nucleic Acids Res.* **51**, e83 (2023).
59. Néron, B. et al. IntegronFinder 2.0: Identification and analysis of integrons across bacteria, with a focus on antibiotic resistance in Klebsiella. *Microorganisms* **10**, 700 (2022).
60. Chen, L., Xiong, Z., Sun, L., Yang, J. & Jin, Q. VFDB 2012 update: Toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res.* **40**, D641–D645 (2012).
61. Starikova, E. V. et al. Phigaro: High-throughput prophage sequence annotation. *Bioinformatics* **36**, 3882–3884 (2020).

Author contributions

Ei Phway Thant: Conceptualization; Methodology; Formal analysis; Investigation; Data curation; Writing – original draft preparation. Chollachai Klaysubun: Methodology; Validation; Visualization; Writing – review and editing. Prasit Palittapongarnpim: Resources; Supervision; Writing – review and editing. Kamonnut Singkhamanan: Methodology; Formal analysis; Investigation; Visualization. Mingkwan Yingkajorn: Sample collection; Clinical data acquisition; Validation. Thunchanok Yaikhan: Methodology; Data curation; Writing – review and editing. Nattarika Chaichana: Laboratory experiments; Data curation. Rattanaruji Pomwised: Bioinformatics analysis; Visualization. Monwadee Wonglapsuwan: Bioinformatics analysis; Validation. Sarunyou Chusri: Clinical interpretation; Writing – review and editing. Arnon Chukamnerd: Clinical interpretation; Supervision. Sissades Tongsima: Resources; Supervision; Writing – review and editing. Worawich Phornsiricharoenphant: Bioinformatics support; Software. Pimonwan Phokhaphan: Bioinformatics support; Visualization. Sirinthorn Sunthornthummas: Data analysis; Writing – review and editing. Nirinya Sudtachat: Resources; review and editing. Chumpol Ngamphiw: Bioinformatics supervision; Validation. Komwit Surachat: Conceptualization; Funding acquisition; Project administration; Supervision; Writing – review and editing.

Funding

This study was supported by the Faculty of Medicine, Prince of Songkla University, and the National Science and Technology Development Agency (NSTDA) through the Health and Medical Research Consortium (Grant No. JRA-CO-2566-19219-TH). Additional financial support was provided by the PSU-Med ISG Scholarship, Faculty of Medicine, Prince of Songkla University, and the Thesis Research Grant for Fiscal Year 2024 from the Graduate School, Prince of Songkla University, Songkhla, Thailand.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Human Research Ethics Committee of Prince of Songkla University (REC65-302–14-1, date of approval: 8 August 2022). Due to the retrospective nature of the study, the Human Research Ethics Committee of Prince of Songkla University waived the need of obtaining informed consent.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-47022-7.

Correspondence and requests for materials should be addressed to K.S.